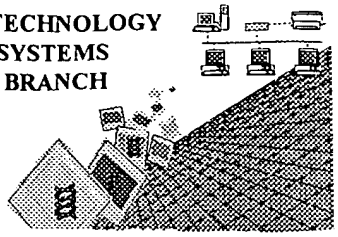
BIOTECHNOLOGY
SYSTEMS
BRANCH



RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/049,280A
Source: PCT/10
Date Processed by STIC: 6/12/2002

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.
PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

1) **INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,**
2) **TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY**

FOR CRF SUBMISSION QUESTIONS, PLEASE CONTACT MARK SPENCER, 703-308-4212.

FOR SEQUENCE RULES INTERPRETATION, PLEASE CONTACT ROBERT WAX, 703-308-4216.
PATENTIN 2.1 e-mail help: patin21help@uspto.gov or phone 703-306-4119 (R. Wax)
PATENTIN 3.0 e-mail help: patin3help@uspto.gov or phone 703-306-4119 (R. Wax)

TO **REDUCE** ERRORED SEQUENCE LISTINGS, **PLEASE** USE THE **CHECKER VERSION 3.1 PROGRAM**, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS:
http://www.uspto.gov/web/offices/pac/checker

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.
Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.
Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<**http://www.uspto.gov/ebc/efs/downloads/documents.htm**>, **EFS Submission User Manual** - ePAVE)
2. **U.S. Postal Service:** U.S. Patent and Trademark Office, Box Sequence, P.O. Box 2327, Arlington, VA 22202
3. **Hand Carry directly to:**
   U.S. Patent and Trademark Office, Technology Center 1600, Reception Area, 7th Floor, Examiner Name, Sequence Information, Crystal Mall One, 1911 South Clark Street, Arlington, VA 22202
   Or
   U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. **Federal Express, United Parcel Service, or other delivery service to:** U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 01/29/2002

Raw Sequence Listing Error Summary

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: 10/049,280A

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE

1	Wrapped Nucleics Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
2	Invalid Line Length	The rules require that a line not exceed 72 characters in length. This includes white spaces.
3	Misaligned Amino Numbering	The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
4	Non-ASCII	The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
5	Variable Length	Sequence(s)____contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
6	PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
7	Skipped Sequences (OLD RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
8	Skipped Sequences (NEW RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence. <210> sequence id number <400> sequence id number 000
9	Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
10	Invalid <213> Response	Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
11	Use of <220>	Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
12	PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
13	Misuse of n	n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.

PCT10

RAW SEQUENCE LISTING DATE: 06/12/2002
PATENT APPLICATION: US/10/049,280A TIME: 14:53:23

Input Set : A:\seq list.txt
Output Set: N:\CRF3\06122002\J049280A.raw

3 <110> APPLICANT: BASF Aktiengesellschaft
5 <120> TITLE OF INVENTION: Homogentisate-Dioxygenase
7 <130> FILE REFERENCE: M/40226
9 <140> CURRENT APPLICATION NUMBER: US/10/049,280A
10 <141> CURRENT FILING DATE: 2002-05-08
12 <160> NUMBER OF SEQ ID NOS: 15
14 <170> SOFTWARE: PatentIn Ver. 2.1
16 <210> SEQ ID NO: 1
17 <211> LENGTH: 575
18 <212> TYPE: DNA
19 <213> ORGANISM: Brassica napus
21 <220> FEATURE:
22 <221> NAME/KEY: misc_feature
23 <222> LOCATION: (1)..(6)
24 <223> OTHER INFORMATION: /function= "restriction site"
26 <220> FEATURE:
27 <221> NAME/KEY: misc_feature
28 <222> LOCATION: (570)..(575)
29 <223> OTHER INFORMATION: /function = "restriction site"
31 <400> SEQUENCE: 1
32 gtcgacgggc cgatgggggc gaagggtctt gctgcaccaa gagattttct tgcaccaacg 60
33 gcatggtttg aggaagggct acggcctgac tacactattg ttcagaagtt tggcggtgaa 120
34 ctctttactg ctaaacaaga tttctctccg ttcaatgtgg ttgcctggca tggcaattac 180
35 gtgccttata agtatgacct gcacaagttc tgtccataca acactgtcct tgtagaccat 240
36 ggagatccat ctgtaaatac agttctgaca gcaccaacgg ataaacctgg tgtggccttg 300
37 cttgattttg tcatattccc tcctcgttgg ttggttgctg agcatacctt tcgacctcct 360
38 tactaccatc gtaactgcat gagtgaattt atgggcctaa tctatggtgc ttacgaggcc 420
39 aaagctgatg gatttctacc tggtggcgca agtcttcaca gttgtatgac acctcatggt 480
40 ccagatacaa ccacatacga ggcgacgatt gctcgtgtaa atgcaatggc tccttataag 540
41 ctcacaggca ccatggcctt catgtttgag gtacc 575
44 <210> SEQ ID NO: 2
45 <211> LENGTH: 26
46 <212> TYPE: DNA
47 <213> ORGANISM: artificial sequence
49 <220> FEATURE:
50 <223> ~~OTHER INFORMATION~~: description of artificial sequence: /desc =
51 "oligonucleotide"
53 <220> FEATURE:
54 <221> NAME/KEY: misc_feature
55 <222> LOCATION: (9)
56 <223> OTHER INFORMATION: /mod_base = i
58 <220> FEATURE:
59 <221> NAME/KEY: misc_feature

Does Not Comply
Corrected Diskette Needed

pp 1-4, 6

insufficient explanation - give source of genetic material

(see item 11 on Error Summary Sheet)

RAW SEQUENCE LISTING DATE: 06/12/2002
PATENT APPLICATION: US/10/049,280A TIME: 14:53:23

Input Set : A:\seq list.txt
Output Set: N:\CRF3\06122002\J049280A.raw

60 <222> LOCATION: (12)
61 <223> OTHER INFORMATION: /mod_base = i
63 <220> FEATURE:
64 <221> NAME/KEY: misc_feature
65 <222> LOCATION: (15)
66 <223> OTHER INFORMATION: /mod_base = i
68 <220> FEATURE:
69 <221> NAME/KEY: misc_feature
70 <222> LOCATION: (18)
71 <223> OTHER INFORMATION: /mod_base = i
73 <220> FEATURE:
74 <221> NAME/KEY: misc_feature
75 <222> LOCATION: (21)
76 <223> OTHER INFORMATION: /mod_base = i
78 <220> FEATURE:
79 <221> NAME/KEY: misc_feature
80 <222> LOCATION: (24)
81 <223> OTHER INFORMATION: /mod_base = i
83 <400> SEQUENCE: 2
W--> 84 **gtcgacggnc cnatnggngc naangg** 26
87 <210> SEQ ID NO: 3
88 <211> LENGTH: 29
89 <212> TYPE: DNA
90 <213> ORGANISM: artificial sequence
92 <220> FEATURE:
93 <223> OTHER INFORMATION: description of artificial sequence: /desc =
94 "oligonucleotide"
96 <220> FEATURE:
97 <221> NAME/KEY: misc_feature
98 <222> LOCATION: (18)
99 <223> OTHER INFORMATION: /mod_base = i
101 <220> FEATURE:
102 <221> NAME/KEY: misc_feature
103 <222> LOCATION: (24)
104 <223> OTHER INFORMATION: /mod_base = i
106 <220> FEATURE:
107 <221> NAME/KEY: misc_feature
108 <222> LOCATION: (27)
109 <223> OTHER INFORMATION: /mod_base = i
111 <400> SEQUENCE: 3
W--> 112 **ggtacctcra acatraangc catngtncc** 29
115 <210> SEQ ID NO: 4
116 <211> LENGTH: 25
117 <212> TYPE: DNA
118 <213> ORGANISM: artificial sequence
120 <220> FEATURE:
121 <223> OTHER INFORMATION: description of artificial sequence: /desc =
122 "oligonucleotide"
124 <400> SEQUENCE: 4

RAW SEQUENCE LISTING DATE: 06/12/2002
PATENT APPLICATION: US/10/049,280A TIME: 14:53:23

Input Set : A:\seq list.txt
Output Set: N:\CRF3\06122002\J049280A.raw

```
125 gaattcgatc tgtcgtctca aactc                                          25
128 <210> SEQ ID NO: 5
129 <211> LENGTH: 26
130 <212> TYPE: DNA
131 <213> ORGANISM: artificial sequence
133 <220> FEATURE:
134 <223> OTHER INFORMATION: description of artificial sequence: /desc =
135       "oligonucleotide"
137 <400> SEQUENCE: 5
138 ggtaccgtga tagtaaacaa ctaatg                                          26
141 <210> SEQ ID NO: 6
142 <211> LENGTH: 34
143 <212> TYPE: DNA
144 <213> ORGANISM: artificial sequence
146 <220> FEATURE:
147 <223> OTHER INFORMATION: description of artificial sequence: /desc =
148       "oligonucleotide"
150 <400> SEQUENCE: 6
151 atggtacctt ttttgcataa acttatcttc atag                                34
154 <210> SEQ ID NO: 7
155 <211> LENGTH: 43
156 <212> TYPE: DNA
157 <213> ORGANISM: artificial sequence
159 <220> FEATURE:
160 <223> OTHER INFORMATION: description of artificial sequence: /desc =
161       "oligonucleotide"
163 <400> SEQUENCE: 7
164 atgtcgaccc gggatccagg gccctgatgg gtcccatttt ccc                       43
167 <210> SEQ ID NO: 8
168 <211> LENGTH: 25
169 <212> TYPE: DNA
170 <213> ORGANISM: artificial sequence
172 <220> FEATURE:
173 <223> OTHER INFORMATION: description of artificial sequence: /desc =
174       "oligonucleotide"
176 <400> SEQUENCE: 8
177 gtcgacgaat ttccccgaat cgttc                                          25
180 <210> SEQ ID NO: 9
181 <211> LENGTH: 24
182 <212> TYPE: DNA
183 <213> ORGANISM: artificial sequence
185 <220> FEATURE:
186 <223> OTHER INFORMATION: description of artificial sequence: /desc =
187       "oligonucleotide"
189 <400> SEQUENCE: 9
190 aagcttccga tctagtaaca taga                                           24
193 <210> SEQ ID NO: 10
194 <211> LENGTH: 25
195 <212> TYPE: DNA
```

RAW SEQUENCE LISTING DATE: 06/12/2002
PATENT APPLICATION: US/10/049,280A TIME: 14:53:23

Input Set : A:\seq list.txt
Output Set: N:\CRF3\06122002\J049280A.raw

196 <213> ORGANISM: artificial sequence
198 <220> FEATURE:
199 <223> OTHER INFORMATION: description of artificial sequence: /desc =
200 "oligonucleotide"
202 <400> SEQUENCE: 10
203 aagcttgatc tgtcgtctca aactc 25
206 <210> SEQ ID NO: 11
207 <211> LENGTH: 24
208 <212> TYPE: DNA
209 <213> ORGANISM: artificial sequence
211 <220> FEATURE:
212 <223> OTHER INFORMATION: description of artificial sequence: /desc =
213 "oligonucleotide"
215 <400> SEQUENCE: 11
216 aagcttccga tctagtaaca taga 24
219 <210> SEQ ID NO: 12
220 <211> LENGTH: 32
221 <212> TYPE: DNA
222 <213> ORGANISM: artificial sequence
224 <220> FEATURE:
225 <223> OTHER INFORMATION: description of artificial sequence: /desc =
226 "oligonucleotide"
228 <400> SEQUENCE: 12
229 attctagaca tggagtcaaa gattcaaata ga 32
232 <210> SEQ ID NO: 13
233 <211> LENGTH: 32
234 <212> TYPE: DNA
235 <213> ORGANISM: artificial sequence
237 <220> FEATURE:
238 <223> OTHER INFORMATION: description of artificial sequence: /desc =
239 "oligonucleotide"
241 <400> SEQUENCE: 13
242 attctagagg acaatcagta aattgaacgg ag 32
245 <210> SEQ ID NO: 14
246 <211> LENGTH: 1159
247 <212> TYPE: DNA
248 <213> ORGANISM: artificial sequence
250 <220> FEATURE:
251 <223> OTHER INFORMATION: description of artificial sequence: /desc =
252 "DNA"
254 <220> FEATURE:
255 <221> NAME/KEY: misc_feature
256 <222> LOCATION: (1)..(6)
257 <223> OTHER INFORMATION: /function = "restriction site"
259 <220> FEATURE:
260 <221> NAME/KEY: CDS
261 <222> LOCATION: (8)..(1153)
263 <220> FEATURE:
264 <221> NAME/KEY: misc_feature

RAW SEQUENCE LISTING DATE: 06/12/2002
PATENT APPLICATION: US/10/049,280A TIME: 14:53:23

Input Set : A:\seq list.txt
Output Set: N:\CRF3\06122002\J049280A.raw

```
<222> LOCATION: (1154)..(1159)
<223> OTHER INFORMATION: /function = "restriction site"
<400> SEQUENCE: 14
gtcgact atg act caa act act cat cat act cca gat act gct aga caa   49
        Met Thr Gln Thr Thr His His Thr Pro Asp Thr Ala Arg Gln
          1               5                  10
gct gat cct ttt cca gtt aag gga atg gat gct gtt gtt ttc gct gtt   97
Ala Asp Pro Phe Pro Val Lys Gly Met Asp Ala Val Val Phe Ala Val
 15                  20                  25                  30
gga aac gct aag caa gct gct cat tac tac tct act gct ttc gga atg   145
Gly Asn Ala Lys Gln Ala Ala His Tyr Tyr Ser Thr Ala Phe Gly Met
                 35                  40                  45
caa ctt gtt gct tac tct gga cca gaa aac gga tct aga gaa act gct   193
Gln Leu Val Ala Tyr Ser Gly Pro Glu Asn Gly Ser Arg Glu Thr Ala
             50                  55                  60
tct tac gtt ctt act aac gga tct gct aga ttc gtt ctt act tct gtt   241
Ser Tyr Val Leu Thr Asn Gly Ser Ala Arg Phe Val Leu Thr Ser Val
         65                  70                  75
att aag cca gct acc cca tgg gga cat ttc ctt gct gat cac gtt gct   289
Ile Lys Pro Ala Thr Pro Trp Gly His Phe Leu Ala Asp His Val Ala
     80                  85                  90
gaa cac gga gat gga gtt gtt gat ctt gct att gaa gtt cca gat gct   337
Glu His Gly Asp Gly Val Val Asp Leu Ala Ile Glu Val Pro Asp Ala
 95                 100                 105                 110
aga gct gct cat gct tac gct att gaa cat gga gct aga tct gtt gct   385
Arg Ala Ala His Ala Tyr Ala Ile Glu His Gly Ala Arg Ser Val Ala
                115                 120                 125
gaa cca tac gaa ctt aag gat gaa cat gga act gtt gtt ctt gct gct   433
Glu Pro Tyr Glu Leu Lys Asp Glu His Gly Thr Val Val Leu Ala Ala
            130                 135                 140
att gct act tac gga aag act aga cat act ctt gtt gat aga act gga   481
Ile Ala Thr Tyr Gly Lys Thr Arg His Thr Leu Val Asp Arg Thr Gly
        145                 150                 155
tac gat gga cca tac ctt cca gga tac gtt gct gct gct cca att gtt   529
Tyr Asp Gly Pro Tyr Leu Pro Gly Tyr Val Ala Ala Ala Pro Ile Val
    160                 165                 170
gaa cca cca gct cat aga acc ttc caa gct att gac cat tgt gtt ggt   577
Glu Pro Pro Ala His Arg Thr Phe Gln Ala Ile Asp His Cys Val Gly
175                 180                 185                 190
aac gtt gaa ctc gga aga atg aac gaa tgg gtt gga ttc tac aac aag   625
Asn Val Glu Leu Gly Arg Met Asn Glu Trp Val Gly Phe Tyr Asn Lys
                195                 200                 205
gtt atg gga ttc act aac atg aag gaa ttc gtt gga gat gat att gct   673
Val Met Gly Phe Thr Asn Met Lys Glu Phe Val Gly Asp Asp Ile Ala
            210                 215                 220
act gag tac tct gct ctt atg tct aag gtt gtt gct gat gga act ctt   721
Thr Glu Tyr Ser Ala Leu Met Ser Lys Val Val Ala Asp Gly Thr Leu
        225                 230                 235
aag gtt aaa ttc cca att aat gaa cca gct ctt gct aag aag aag tct   769
```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 06/12/2002
PATENT APPLICATION: **US/10/049,280A** TIME: 14:53:24

Input Set : **A:\seq list.txt**
Output Set: **N:\CRF3\06122002\J049280A.raw**

Please Note:
Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:2; N Pos. 9,12,15,18,21,24
Seq#:3; N Pos. 18,24,27